Fast Entropy-Based Nonrigid Registration*

Eduardo Suárez[1], Jose A. Santana[1], Eduardo Rovaris[1],
Carl-Fredrik Westin[2], and Juan Ruiz-Alzola[1,2]

[1] Medical Technology Center,
Univ. Las Palmas of GC & Gran Canaria
Dr. Negrín Hospital, SPAIN
eduardo@ctm.ulpgc.es
[2] Laboratory of Mathematics in Imaging,
Brigham and Women's Hospital and
Harvard Medical School, USA

Abstract. Computer vision tasks such as learning, recognition, classification or segmentation applied to spatial data often requires spatial normalization of repeated features and structures. Spatial normalization, or in other words, image registration, is still a big hurdle for the image processing community. Its formulation often relies on the fact that correspondence is achieved when a similarity measure is maximized. This paper presents a novel similarity measuring technique based on a *coupling function* inside a template matching framework. It allows using any entropy-based similarity metric, which is crucial for registration using different acquisition devices. Results are presented using this technique on a multiresolution incremental scheme.

1 Introduction

Registration consists of finding the spatial correspondence between two coordinate systems with a scalar field defined on each one. For the two and three dimensional case, this is commonly called image registration. The correspondence is satisfied by means of a known similarity in the two datasets. This similarity is dependent on the application and is defined by high level information such as geodesic points and textures in satellite images, relevant anatomical points in medical images, or any other features for stereo matching or biosensing.

Datasets to be registered can therefore correspond to the same or to different subjects in the case of medical imaging or recognition, and can also be from the same or from different imaging modalities such as the different channels in satellite sensing. Putting into correspondence two images that can be topologically different (for example, in the case of a medical pathology) and where the pixel intensities measure different physical magnitudes (multimodality) poses a serious challenge that has sparked intensive research over the last years [1].

* This work was supported by the spanish Ministry of Science and Technology and European Commission, co-funded grant TIC-2001-38008-C02-01, NIH grant P41-RR13218 and CIMIT, and spanish FPU grant AP98-52835909.

R. Moreno-Díaz and F. Pichler (Eds.): EUROCAST 2003, LNCS 2809, pp. 607–615, 2003.

Particularly, in brain imaging, nonrigid registration is crucial for spatial normalization and brain understanding. Numerical approaches to nonrigid registration often relies on regularization theory, which separates the similarity measures of the image features from the smoothness constraints of the warping.

Voxel based registration methods can broadly be divided in to two categories: methods based on a template matching technique, and those based on a variational approach. Template matching, also known as block matching in MPEG compression, finds the displacement for every voxel in a source image by maximizing a local similarity measure, obtained from a small neighborhood of the source image and a set of potential correspondent neighborhoods in a target image. The main disadvantages of its conventional formulation are that

- it estimates the displacement field independently in every voxel and thus spatial coherence is imposed to the solution
- it needs to test several discrete displacements to find a minimum and
- it has the inability of making a good match when no discriminant structure is available, such as in homogeneous regions, surfaces and edges (aperture problem).

Template matching was popular years ago due to its conceptual simplicity [2], but it does not impose any constraint on the resulting discrete fields, loosing its place in this arena.

On the other hand, variational methods rely on the minimization of a functional (energy) that is usually formulated as the addition of two terms: data coupling and regularization, the former forcing the similarity between both datasets (target, and source deformed with the estimated field) to be high while the later forcing the estimated field to fulfill some constraint (usually enforcing spatial coherence-smoothness).

Studies have already shown the power of the entropy-based similarity measures [3] to deal with multimodal registration. Their main drawback for nonrigid registration is the joint probability density function estimation, which must be known for every voxel displacement and computed on small neighborhoods of such voxels. There are two main approaches to overcome this problem. The first one [4], is to consider a global functional of a parameterized warping, and perform optimization on the parameters. It needs successive interpolation and global pdf estimation on every step, increasing the computational cost. The second approach is to compute local estimations of the gradient of the probability density function (hereinafter pdf) in order to steer local optimizations. This approach has been used in [5,6]. Nevertheless, they use a global pdf that is taken as static in gradient estimation.

In this paper we propose an entropy-based similarity measurement technique which can be used in a registration algorithm. It computes local entropy-based similarities, while using the appropriate probability density functions.

In order to show the efficiency of the matching technique a registration framework is needed. Section 2 introduces the registration framework exposed in [7] for later use. Section 3 describes the similarity measurement technique. Results are shown in Section 4 and conclusions in Section 5.

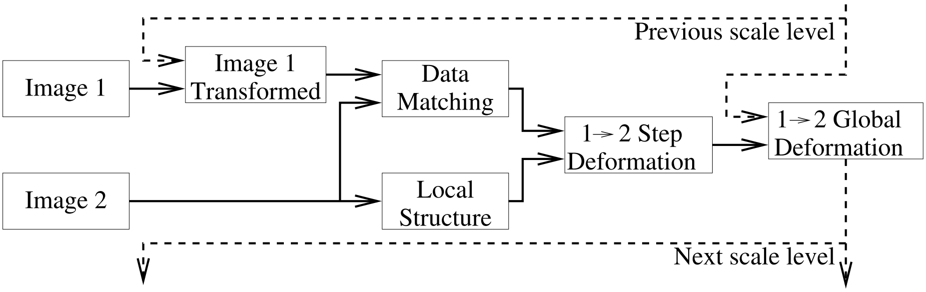


Fig. 1. Algorithm pipeline for one pyramidal level.

2 Nonrigid Registration

In order to introduce the entropy estimation scheme, it is first necessary to describe the registration framework where it is going to be used.

2.1 Multiscale

The registration algorithm consists of a pyramidal block-matching scheme, where the deformation field is regularized by weighting with local structure [7]. The algorithm works similarly to Kovačič and Bajcsy incremental multiresolution matching [8], which is based on a gaussian multiscale pyramidal representation. In the highest level, the deformation field is estimated by regularized template matching steered by the local structure of the image (details in sections below). In the next level, the source dataset is deformed with a deformation field obtained by spatial interpolation of the one obtained in the previous level. The deformed source and the target datasets on the current level are then registered to obtain the deformation field corresponding to the current level of resolution. This process is propagated to every level in the pyramid. The algorithm implementation is summarized in figure 1.

2.2 Data Matching

Template matching finds the displacement for every voxel in a source image by minimizing a local cost measure, obtained from a small neighborhood of the source image and a set of potential correspondent neighborhoods in a target image. The main disadvantage of template matching is that it estimates the displacement field independently in every voxel and no spatial coherence is imposed to the solution. Another disadvantage of template matching is that it needs to test several discrete displacements to find a minimum.

There exists some optimization-based template matching solutions that provide a real solution for every voxel, though they are slow [9]. Therefore, most template matching approaches render discrete displacement fields. Another problem associated to template matching is commonly denoted as the *aperture problem* in the computer vision literature [10]. This essentially consists of the inability of making a good match when no discriminant structure is available, such as in homogeneous regions, surfaces and edges. When this fact is not taken into account the matching process is steered by noise and not by the local structure, since it is not available.

The registration framework keeps the simplicity of template matching while it addresses its drawbacks. Indeed the algorithm presented here consists of a weighted regularization of the template matching solution, where weights are obtained from the local structure, in order to render spatially coherent real deformation fields. Thanks to the multiscale nature of our approach only displacements of one voxel are necessary when matching the local neighborhoods.

2.3 Local Structure

Local structure measures the quantity of discriminant spatial information on every point of an image and it is crucial for template matching performance: the higher the local structure, the better the result obtained on that region with template matching.

In order to quantify local structure, a structure tensor is defined as $\mathbf{T}(\mathbf{x}) = (\nabla I(\mathbf{x}) \cdot \nabla I(\mathbf{x})^t)_\sigma$, where the subscript σ indicates a local smoothing. The structure tensor consists of a symmetric positive-semidefinite $D \times D$ matrix that can be associated to ellipses, i.e., eigenvectors and eigenvalues correspond to the ellipses axes directions and lengths respectively. D denotes the dimension of the space. A scalar measure of the local structure can be obtained as [11,12].

$$structure(\mathbf{x}) = \frac{\left(\det \mathbf{T}(\mathbf{x})\right)^{2/D}}{\text{trace}\, \mathbf{T}(\mathbf{x})}. \quad (1)$$

Small eigenvalues indicate the lack of gradient variation along the associated principal direction and, therefore, high structure would be represented by big rounded (no eigenvalue is small) ellipses. This way, anatomical landmarks will have the highest measure of local structure. Curves will be detected with lower intensity than points and surfaces will have even lower intensity. Homogeneous areas have almost no structure.

2.4 Spatial Regularization

The nature of the problem imposes the spatial mapping to be a diffeomorphism, therefore the Jacobian of the deformation field must be positive. This condition is preserved while regularizing small deformation fields. Large deformation fields with no spatial coherence would need a very destructive regularization to achieve invertibility. Thus, only one pixel displacements are allowed for matching. For every level of the pyramid the mapping is obtained by composing the transformation on the higher level with the one on the current level, in order to preserve the Jacobian positiveness condition.

Regularization is achieved by means of the *Normalized Convolution* [13], a refinement of weighted-least squares that explicitly deals with the so-called signal/certainty philosophy. Essentially the scalar measure of structure (the certainty) is incorporated as a weighting function in a least squares fashion. The field (the signal) obtained from template matching is then projected onto a vector space described by a non-orthogonal basis, i.e., the dot products between the field and every element of the basis provide co-variant components that must be converted into contravariant by an appropriate metric tensor. Normalized convolution provides a simple and efficient implementation of this operation.

Moreover, an applicability function is enforced on the basis elements in order to guarantee a proper localization and to avoid high frequency artifacts. This essentially corresponds weighting each basis element with the applicability function. In this application a Gaussian function will be used. Convolution with a Gaussian window can be implemented in a fast and efficient way, because it is a separable kernel.

In the three-dimensional case, we have used a basis consisting in three elements, which can be written as:

$$\mathbf{b}_1 \begin{cases} d_1(\mathbf{x}) = 1 \\ d_2(\mathbf{x}) = 0 \\ d_3(\mathbf{x}) = 0 \end{cases} \quad \mathbf{b}_2 \begin{cases} d_1(\mathbf{x}) = 0 \\ d_2(\mathbf{x}) = 1 \\ d_3(\mathbf{x}) = 0 \end{cases} \quad \mathbf{b}_3 \begin{cases} d_1(\mathbf{x}) = 0 \\ d_2(\mathbf{x}) = 0 \\ d_3(\mathbf{x}) = 1 \end{cases} \tag{2}$$

This way, we decouple the components of the displacement field and regularization is done applying independently normalized convolution to each component.

3 Entropy Estimation

In [7], the registration framework was tested using square blocks that were matched using the sum of squared differences and correlation coefficient as similarity measures. In this work we introduce entropy-based similarity measures into this framework, though it can be used by any algorithm based on template matching.

A similarity measure can be interpreted as a function defined on the joint probability space of two random variables to be matched. In the case of block matching, each block represents a set of samples from each random variable.

When this probability density function is known, mutual information can be computed as

$$MI(I_1, I_2) = \int_{\Omega} p(i_1, i_2) \log \frac{p(i_1, i_2)}{p(i_1)p(i_2)} \, di_1 \, di_2 \tag{3}$$

where I_1, I_2 are the images to register and Ω is the joint probability function space.

A discrete approximation is to compute the mutual information from the pdf and a small number N of samples $(i_1[k], i_2[k])$:

$$MI(I_1, I_2) \backsimeq \sum_{k=1}^{N} \log \frac{p(i_1[k], i_2[k])}{p(i_1[k])p(i_2[k])} = \sum_{k=1}^{N} f_p(i_1[k], i_2[k]) \ , \tag{4}$$

where f_p is a *coupling function* defined on Ω.

Therefore, the local evaluation of the mutual information for a displaced block containing N voxels can be computed just by summing the coupling function f_p on the k samples that belong to this block.

We propose to compute a set of multidimensional images, each of them containing at each voxel the local similarity measure corresponding to a single displacement applied

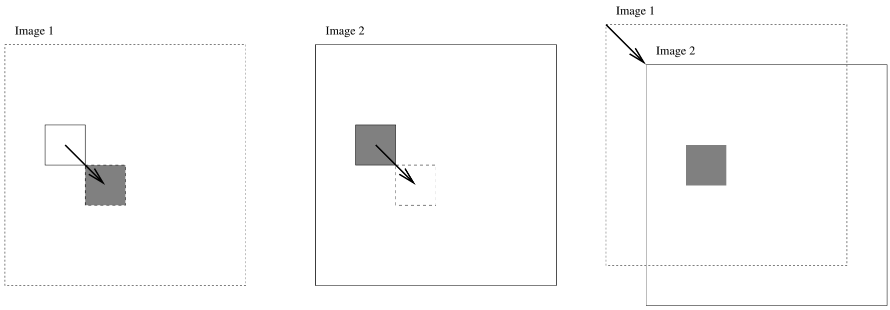


Fig. 2. *Left*: Target image to be matched; *Center*: Reference image where similarity measure is going to be estimated for every discrete displacement; *Right*: For every discrete displacement the similarity measure is computed for every voxel by performing a convolution.

to the whole target image. A decision will be made for each voxel depending on which displacement renders the biggest similarity.

A problem associated with local entropy-based similarity measures is the local estimation of the joint pdf of both blocks, since never enough samples are available. We propose to overcome this problem by using the joint pdf corresponding to the whole displaced source image and the target one.

The pdf to be used for a given displacement will be the global joint intensity histogram of the reference image with the displaced target image. This is crucial for higher pyramidal levels, where one voxel displacement changes drastically the pdf estimation.

It is straightforward to compute the local mutual information for a given discrete displacement in the whole image. This requires only the convolution of a square kernel representing the block window and the evaluation of the coupling function for every pair of voxels. Furthermore, since the registration framework only needs discrete deformation fields, no interpolation is needed in this step. Any similarity measure which can be computed as a kernel convolution can be implemented this way. A small sketch of this technique is shown in figure 2. For smoothness and locality reasons, we have chosen to convolve using gaussian kernels instead of square ones.

In order to achieve a further computational saving, equation 4 can be written as:

$$MI(I_1, I_2) \backsimeq \sum_{k=1}^{N} \big(\log p(i_1[k], i_2[k]) - \log p(i_1[k]) - \log p(i_2[k]) \big) \quad (5)$$

Since the source image remains static, the term $\log p(i_1[k])$ is the same across the whole set of images with the local mutual information for each displacement. Therefore, this term does not affect to the decision rule, and can be disregarded, hence, avoiding to estimate the marginal pdf for the target image.

The displacement field tells about the displacement of a voxel in the source image. The similarity measure will be referred to the source image reference system (image 1). For a given voxel in the source image, the comparison of equation 5 for different displacement will always contain the same terms depending on $p(i_1[k])$. So that, we can take this term off and modify accordingly the coupling function to save computation cost.

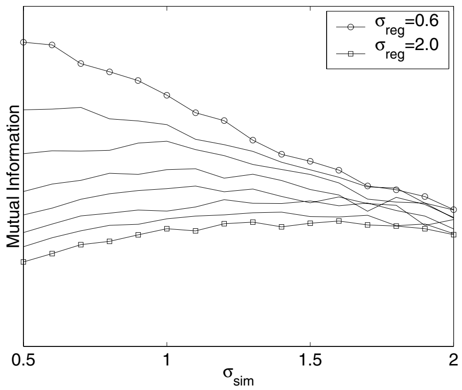


Fig. 3. Mutual Information as a function of the standard deviations used for similarity and regularization kernels.

Any other entropy-based similarity measure can be estimated in a similar way. The computational cost is then very similar to any other similarity measure not based on entropy

4 Results

Two experiments have been carried out. The former will give us quantitative results about the registration algorithm, while the latter will show us qualitative ones. Mutual information is the selected similarity measure used for next registration results. Datasets have previously been rigidly aligned and resampled.

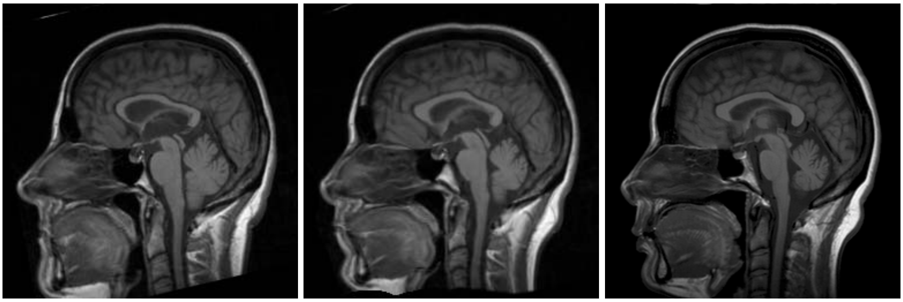
In the first experiment, two T1-weighted images corresponding to different patients have been registered. The datasets have $90 \times 98 \times 72$ voxels. Seven multiscale pyramidal levels have been used.

To carry out the experiment, two kernel standard deviations σ_{sim} and σ_{reg} are needed. The former will used for the convolution kernel of the coupling function sum in equation 4. The latter is for the kernel used in the regularization of the displacement field, that is, the normalized convolution applicability. Figure 3 shows the mutual information of the registered images as a function of σ_{sim} and σ_{reg}.

In a second experiment, an interpatient nonrigid registration has been done with two T1-weighted images of size 256×256, with $\sigma_{reg} = 1.2$ and $\sigma_{sim} = 1.2$ and seven pyramidal levels. Registration is shown in figure 4. We can see the source image deformed onto the target image.

5 Conclusions and Future Work

Figure 3 shows that we get higher values of mutual information when using small σ_{reg} and σ_{sim}. However, by visual inspection, we have checked that the deformation is not smooth at all, and hence results are not valid using these values. As we move to very high σ_{reg} values, the regularization is so strong that it makes the deformation to approach zero, making similarity drop to the similarity of the unregistered images.



(a) Source (b) Registered (c) Target

Fig. 4. Interpatient registration.

For middle values of σ_{reg} we find a reasonable behaviour in the local estimation of the similarity measure, that is, for high σ_{sim} the registration is worse because the measure is not local anymore, and for low values, there are to few samples to have a good estimation of the mutual information. We conclude that local estimation of mutual information has been done successfully with this technique. Visual inspection of the registration in figure 3 confirms this conclusion.

Despite this technique can be applied for intersubject multimodal registration, local variation of intensities may be present in one modality while not in the other, making the registration process fail. So then, its optimum application is for monomodal image registration. There, different scanning parameters, bias, or anatomical differences may occur and classical monomodal measures are not appropriate.

As future work, speedup of the algorithm will be achieved by means of a full C implementation. At the moment, the algorithm is implemented in Matlab and it takes about fifteen minutes for every registration described in experiment 1.

References

1. J.B.A. Maintz and M.A. Viergever. A survey of medical image registration. *Medical Image Analysis*, 2(1):1–36, April 1998.
2. Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis*. John Wiley & Sons, 1973.
3. M. Holden, D.L.G. Hill, E.R.E Denton, J.M. Jarosz, T.C.S. Cox, J. Goody, T. Rohlfing, and D.J. Hawkes. Voxel similarity measures for 3D serial MR image registration. *IEEE Trans. Med. Imag.*, 19(2):94–102, 2000.
4. D. Rueckert, L. I. Sonoda, E.R.E. Denton, S. Rankin, C. Hayes, D. L. G. Hill, M. Leach, and D. J. Hawkes. Comparison and evaluation of rigid and non-rigid registration of breast MR images. In *SPIE Med. Im. 1999: Im. Proc.*, pages 78–88, San Diego, CA, February 1999.
5. N. Hata, T. Dohi, S. K. Warfield, W. M. Wells, R. Kikinis, and F. A. Jolesz. Multimodality deformable registration of pre- and intraoperative images for MRI-guided brain surgery. In *MICCAI '98*, pages 1067–1074, 1998.
6. G. Hermosillo, C. Chefd'Hotel, and O. Faugeras. A variational approach to multi-modal image matching. Technical Report 4117, INRIA, 2001.

7. Eduardo Suárez, C.-F. Westin, E. Rovaris, and Juan Ruiz-Alzola. Nonrigid registration using regularized matching weighted by local structure. In *MICCAI '02*, number 2 in Lecture Notes in Computer Science, pages 581–589, Tokyo, Japan, September 2002.
8. Stanislav Kovacic and R.K. Bajcsy. *Brain Warping*, chapter Multiscale/Multiresolution Representations, pages 45–65. Academic Press, 1999.
9. Eduardo Suárez, Rubén Cárdenes, Carlos Alberola, Carl-Fredrik Westin, and Juan Ruiz-Alzola. A general approach to nonrigid registration: Decoupled optimization. In *23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '01)*. IEEE Engineering in Medicine and Biology Society, october 2001.
10. T. Poggio, V. Torre, and C. Koch. Computational vision and regularization theory. *Nature*, 317:314–319, September 1985.
11. K. Rohr. Differential operators for detecting point landmarks. *Image and Vision Computing*, 15:219–233, 1997.
12. J. Ruiz-Alzola, R. Kikinis, and C.-F. Westin. Detection of point landmarks in multidimensional tensor data. *Signal Processing*, 81:2243–2247, 2001.
13. C-F. Westin. *A Tensor Framework for Multidimensional Signal Processing*. PhD thesis, Linköping University, Sweden, SE-581 83 Linköping, Sweden, 1994. Dissertation No 348, ISBN 91-7871-421-4.
14. Jean-Philippe Thiran and Torsten Butz. Fast non-rigid resistration and model-based segmentation of 3d images using mutual information. In *SPIE*, volume 3979, pages 1504–1515, San Diego, USA, 2000.
15. P. Thévenaz and M. Unser. Optmization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12):2083–2099, 2000.
16. F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens. Multimodality image registration by maximization of mutual information. *IEEE Transactions on Medical Imaging*, 16:187–198, 1997.
17. J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. A multiscale approach to mutual information matching. In K.M. Hanson, editor, *Medical Imaging: Image Processing*, volume 3338 of *Proc. SPIE*, pages 1334–1344, Bellingham, WA, 1998. SPIE Press.
18. J. B. Antoine Maintz, Erik H. W. Meijering, and Max A. Viergever. General multimodal elastic registration based on mutual information. In K. M. Hanson, editor, *Medical Imaging 1998: Image Processing*, volume 3338 of *Proceedings of SPIE*, pages 144–154, Bellingham, WA, 1998. The International Society for Optical Engineering.
19. J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever. Image registration by maximization of combined mutual information and gradient information. *IEEE Transactions on Medical Imaging*, 19(8):809–814, 2000.
20. William M. Wells, Paul Viola, Hideki Atsumi, Shin Nakajima, and Ron Kikinis. Multi-modal volume registration by maximization of mutual information. *Medical Image Analysis*, 1:35–52, 1996.
21. Paul Viola and Williams M. Wells. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24:137–154, 1997.
22. C. Studholme, D.L.G. Hill, and D.J. Hawkes. Automated 3d registration of mr and ct images of the head. *Medical Image Analysis*, 1(2):163–175, 1996.